

PSYCHIATRIC DIAGNOSIS IN TORONTO, NEW YORK AND LONDON*

G.S. DUCKWORTH, M.D.¹

H.B. KEDWARD, M.D.²

M.R. EASTWOOD, M.D.³

F. ALLODI, M.D.⁴

C. WOOGH, M.D.⁵

R. MEIER, M.D.⁶

"We shall never be able to believe that psychiatry will make a step forward until we decide to throw overboard the whole business of classifications . . . There is but one type of mental disturbance, and we call it insanity" (1). So said Neumann in 1859. There are some who are still saying it today. To Neumann it made little difference what label was attached to a patient since there existed no satisfactory treatment at that time for any major mental illness. Even fifty years later the great Kraepelin was treating schizophrenia with belladonna alkaloids (2). In contrast, the modern clinician has been armed with a formidable cornucopia of potent drugs, each of which possesses definite indications and contraindications. Diagnosis therefore is becoming as critical in psychiatry as in medicine.

National Statistics

National statistics of mental illness suggest great differences in morbidity patterns between Canada, the United States and the United Kingdom (3) — see Table I. The first admission rates for all psychiatric illnesses are higher in Canada than in either the US or the UK. Among diagnostic

groups, the differences between the three centres are highly significant ($p < .0005$), Canada and the U.S. having higher rates of schizophrenia and diseases of the senium, but lower rates of affective illnesses than the UK. The Canadian figures tend to parallel those of the US at a higher level, but do not differ significantly from them. Thus the main differences in morbidity lie between North America and the UK. The US-UK study of psychiatric diagnosis in New York and London established that the divergent rates in those two cities were due entirely to differing diagnostic habits (4). No comparable study has to date been done in Canada.

Three questions immediately suggest themselves: Are the different levels of morbidity indicated by Canadian, US and UK national statistics genuine, or are they the consequence of differing diagnostic habits in the three countries? If so, then are there consistent biases in Canadian diagnoses and is it possible to improve the reliability of diagnoses assigned to patients by the use of standardized techniques? This paper is an attempt to answer some of the questions posed.

Method

This paper stems from two collaborative research projects —

The US-UK cross-national study of psychogeriatric diagnosis (5) which was extended to include Toronto (6), and the Toronto study of chronic psychiatric patients (7).

*This work was performed under grants from the Ontario Mental Health Foundation and Ontario Health Council D.M. 105.

Manuscript received December 1975. Revised March 1976.

¹⁻⁶Psychiatric Epidemiology Unit, Clarke Institute of Psychiatry and Department of Psychiatry, University of Toronto, Toronto, Ontario.

Can. Psychiatr. Assoc. J. Vol. 21 (1976)

NATIONAL STATISTICS

TABLE I
First Admission Rates by Diagnosis, Mental Hospitals for the U.S., England, Wales, and Canada

Diagnosis 1960	Age — Adjusted Rates Per 100,000 Population		
	U.S.A. 1960	England, Wales 1960	Canada
All disorders	102.1	115.7	143.4
I. Schizophrenia	24.7	17.4	25.5
II. Major Affective Disorders	11.0	38.5	17.9
Manic-Depressive	3.3	35.7	11.95
Psychotic Depressive			
Reaction	2.9	DNU	DNU
Involutional Psychotic			
Reaction	4.8	2.8	5.95
III. Diseases of Senium	16.4	12.9	15.0
Cerebral Arteriosclerosis	10.9	1.8	—
Senile Psychosis	5.4	11.1	—
IV. Psychoneurosis	12.4	22.9	31.7
Psychoneurotic Depressive			
Reaction	NA	9.1	17.0
Other Psychoneuroses	NA	13.8	14.7
V. All other Disorders	37.7	24.0	53.3

$X^2 = 33.73$ DF 8
Sig $P < .0005$

U. S. vs England, Wales $X^2 = 22.38$ D.f. = 4 Sig $< .005$
England, Wales vs Canada $X^2 = 18.99$ D.F. = 4 Sig $< .001$
U. S. vs Canada $X^2 = 6.10$ D.F. = 4 N.S.

The Psychogeriatric Sample

The Toronto sample consisted of 100 randomly selected patients admitted to three Toronto hospitals — a provincial hospital (Queen St. Mental Health Centre), a community psychiatric hospital (the Clarke Institute of Psychiatry) and the psychiatric unit of a general hospital (the Toronto General Hospital). All patients fulfilled three criteria for admission to the study: age 65 or over; address within the borough of the city of Toronto; and interviewed within seventy-two hours of admission. Thus, the sample consisted of acute illnesses in elderly patients.

The Chronic Patient Sample (CPS)

The 155 patients in this study were between the ages of 18 and 65 and had been in a provincial mental hospital for two years or longer. They therefore represented a group of young and middle-aged adults suffering from chronic or recurring illnesses.

The Instruments Used

The psychiatrists in the chronic patient study administered a standardized interview, the Present State Examination (PSE), developed by

Wing (8). However, due to the presence in the study of patients suffering from organic brain damage, an additional section was devised by a project psychiatrist (GSD) to aid in the diagnosis of dementia. The PSE was repeated at three-monthly intervals over the succeeding year.

The psychogeriatric study also utilized a standardized technique of interview based on the Geriatric Mental State (GMS) examination which was derived from the PSE of Wing and the Mental Status Schedule (MSS) of Spitzer (9) with additional sections to test the presence of organic dementia. Each patient was interviewed by the project psychiatrists within 72 hours of admission, then after one month and three months. Those receiving a diagnosis of dementia had a retest at one week.

Reliability

Separate reliability studies were conducted for both studies:

GMS — Simultaneous interviews were conducted between the Canadian psychiatrist (GSD) and two US psychiatrists from the US-UK project. Eight interviews were conducted with one interviewer and one observer, and three with one interviewer and two observers. The

psychiatrists rotated alternately from interviewer to observer.

Results of comparison of items rated showed a mean agreement of 92.9 percent on the 8,880 items rated. Other indices also revealed a satisfactory level of agreement (D.I.=0.54, $r=.59$, kappa=.58).

PSE — Ten patient interviews were utilized, each of which was rated by five psychiatrists. Three of these rotated as interviewer and observers in simultaneous interviews, while the remaining two rated videotapes of the interviews. Results of the comparison of the 7,750 items rated showed a mean agreement of 91.6 percent. Other indices also revealed a high level of agreement. (D.I.=.309, $r=.765$, kappa=.762).

The Project Diagnoses

In the CPS study, the project psychiatrist interviewed the patient and rated his responses to the PSE. After the interview was completed, he was given a condensed vignette of the patient's past medical and psychiatric history, prepared in a standardized way. The interviewer then made his provisional diagnosis. The patient was then re-interviewed at three-monthly intervals. The results of the initial examinations were presented to the other project psychiatrists by the interviewer who recounted the facts after deleting the diagnostic label. A consensus diagnosis was then made after discussion of the case by a majority vote of the project psychiatrists, using the International Classification (ICDA) — 8th edition as defined by the WHO glossary. Results of the reliability study showed that there was concordance on 3-digit diagnosis of 98 percent (49/50) and on 4-digit diagnosis of 80 percent. This was very high agreement, and reliability was "spot-checked" during the study to ensure that the high reliability was maintained.

In the psychogeriatric sample study the instrument used was the GMS examination initially and after one and three months. The diagnosis was made by one project psychiatrist, who later presented the data to two other project psychiatrists who made their own diagnoses.

In addition to diagnosis by the doctors, the CPS study also yielded a computer diagnosis by the CATEGO program, based on each separate PSE interview. This was then available for comparison with the doctors' diagnosis and with itself since each patient was re-interviewed up to three times. The only category excluded by CATEGO were the organic patients, who could not be handled by the CATEGO program.

Hospital Diagnoses

Hospital diagnoses were those reached routinely by the attending staff psychiatrist on each patient, following the ICDA-8th ed. This diagnosis was then available for comparison with that reached by the project doctors.

Results

The psychogeriatric study showed that in the sample of 100 the concordance between the hospital and the project diagnoses in the four categories was 75 percent, a higher agreement than the psychogeriatric diagnoses in either the US or the UK.

The chronic patient study showed a concordance between hospital and project diagnoses of 129/155 or 83.23 percent.

Diagnostic trends in Toronto may be delineated by pooling the two samples. The combined sample of 255 comprises patients in age from 18 years upwards, suffering from both acute and chronic illnesses. Thus, it includes the whole spectrum of psychiatric illness except for children's diseases. The agreement between the hospital and project diagnoses in the three centres can be compared. Table II shows that there is great discrepancy between diagnoses rates of schizophrenia in New York, the hospitals labeling over 60 percent of the patients with this illness, while the project psychiatrists found that less than 30 percent received a consensus diagnosis of schizophrenia. These differences were highly significant ($p < .0005$ to χ^2 test). However, in the London and Toronto samples differences were small and in neither case did they attain statistical significance. While these figures demonstrate the similarity between hospital and project diagnoses in Toronto and London, the proportion of illnesses fails to measure agreement or disagreement on individual patients. Figure 1 is a comparison of diagnosis in the Toronto sample. Where both hospital and project doctors agree, the figures follow the shaded diagonal. Disagreement falls outside the shaded area. The figure shows that concordance on diagnoses is present in 204 patients out of a total of 255, an 80 percent level of agreement. This is very high and may be compared with the concordances in

TABLE II
THE HOSPITAL AND PROJECT DIAGNOSES IN THE THREE COUNTRIES

	NEW YORK				LONDON			
	Project (E) Diagnoses		Hospital (O) Diagnoses		Project (E) Diagnoses		Hospital (O) Diagnoses	
	No.	%	No.	%	No.	%	No.	%
Organic	5	(2.6)	10	(5.2)	6	(3.4)	3	(1.7)
Schizophrenia	56	(29.2)	118	(61.5)	61	(35.1)	59	(33.9)
Affective	62	(32.3)	13	(6.8)	76	(43.7)	68	(39.0)
Other	100	(36.0)	82	(26.5)	31	(17.8)	44	(25.2)
Total	223		223		174		174	
				$X^2 = 57.5522$ DF = 3				
				SIG. P < .0005				$X^2 = 3.7311$ Df = 3
								N.S.

	TORONTO			
	Project (E) Diagnoses		Hospital (O) Diagnoses	
	No.	%	No.	%
Organic	75	(29.4)	75	(29.4)
Schizophrenia	101	(39.6)	94	(36.9)
Affective	46	(18.0)	51	(20.0)
Other	33	(12.9)	35	(13.7)
Total	255		255	
				$X^2 = 0.5678$ DF = 3
				N.S.

the US and UK. Concordance on diagnosis is present in 58.74 percent of the US sample (barely over half) and 63.8 percent in the

UK sample. The concordance is highest in Toronto, lowest in New York and in between in London. Thus Toronto physicians, at any rate in the three hospitals studied, do not share the tendency of their colleagues to label most younger patients 'schizophrenic' and most elderly folk 'demented'. Instead, the tendency in Toronto is towards more rigorous use of the diagnostic labels.

Discussion and Conclusions

The present psychogeriatric study was designed to be comparable in methodology with those in New York and London. The findings (6) were that there was no statistically significant difference between the project diagnoses in the three centres. The hospital diagnoses in New York contained an excess of organic dementias not found in the project diagnoses, but, in Toronto, there were no significant differences between the hospital and project diagnoses. However, this study was re-

HOSPITAL AND PROJECT DIAGNOSES IN
TORONTO - POOLED DATA

Project Diagnosis	Hospital Diagnosis				TOTAL
	ORGANIC	SCHIZOPHRENIC	AFFECTIVE	MISCELLANEOUS	
ORGANIC	64	7	2	2	75
SCHIZOPHRENIC	3	81	9	8	101
AFFECTIVE	4	4	36	2	46
MISCELLANEOUS	4	2	4	23	33
TOTAL	75	94	51	35	255

$X^2 = 383.0880$ DF = 9
Sig. P < .0005
Concordance = $\frac{204}{255} = 80.00\%$

Figure 1.

stricted to patients over 65 years of age, and therefore does not provide a representative sample of all Toronto patients. The use of the PSE in the study of chronic psychiatric patients provides reliable diagnostic information for 155 Toronto patients between the ages of 18 and 65. Combining the data of the two studies yields a sample of 255 Toronto patients of all ages and phases of illness examined by five project psychiatrists in three hospitals. This allows ample scope for discrepancies between the diagnoses of project and hospital doctors. The results of this disagreement illuminate diagnostic trends in Toronto which can be contrasted with the trends in New York and London.

Table II demonstrates that of the three, Toronto doctors agreed most often with the consensus diagnoses of the project psychiatrists, and New York doctors least. This finding is unexpected in the light of the US-UK project's videotape comparison of diagnosis in Canada, which found that Canadian doctors tended to take a position closer to their American than their British colleagues (10). In particular, it contradicts the statement that Canadian psychiatrists use the extended concept of schizophrenia which is typical of American doctors.

One possible objection to the findings presented is that chronic patients tend to be a more homogeneous sample than acute patients. The expected improvement of young adults suffering from neuroses and affective disorders would leave a sample lacking these categories but with increased numbers of schizophrenics and organic patients. An estimate of the magnitude of this effect may be obtained by comparing concordance rates of the chronic patient study with the psychogeriatric study. The latter contained 44.0 percent of first admissions and 56.0 percent of readmissions. This proportion does not differ statistically from the ratio of first admissions to readmissions of those over 65 years in all Canada for 1971. Thus, the psychogeriatric sample contains as many acutely ill patients as the general average and considerably more than the 4.9 percent of first admissions and 0 percent of acute patients in the CPS sample.

The concordance rate of 75 percent in the psychogeriatric study is 5 percent below that of the CPS study, and the difference may well be due to the acuteness of the cases in the PG study. However, this difference is not a significant one, and it is unlikely that a sample of acutely ill patients aged 18-65 would yield radically different figures.

Finally, the purist may still object that the project diagnoses, although they are based on standardized interview techniques with high reliability, do not completely eliminate the subjective element when a diagnosis is derived from rated symptoms. In the chronic patient study the use of the 9th edition of the PSE allowed a comparison between the project psychiatrists' diagnoses, and the computer diagnoses obtained from the PSE by the CATEGO program for all except the 'organic' patients who were not handled by CATEGO. Since each patient had up to four separate PSE interviews, it was possible to examine the stability over time of symptoms and diagnoses in these patients. Figure 2 compares the project diagnoses of the functional patients with their CATEGO diagnoses, where a hierarchy of diagnoses is used with CATEGO (if one or more CATEGO

CHRONIC PSYCHOSIS PROJECT DIAGNOSES VS
ALL CATEGO DIAGNOSES

		Catego Diagnosis			
		SCHIZOPHRENIC	AFFECTIVE	MISCELLANEOUS	TOTAL
Project Diagnosis	SCHIZOPHRENIC	53	12	11	76
	AFFECTIVE	0	14	0	14
	MISCELLANEOUS	0	1	6	7
	TOTAL	53	27	17	97

(73)

CONCORDANCE = $73/97 = 75.3\%$

$\chi^2 = 66.69$, DF = 4

Sig. $P < .0005$

Figure 2.

diagnoses is 'schizophrenic' the patient is schizophrenic. If not, and one or more diagnoses is 'affective' the patient is affective. If not, then the diagnosis is 'miscellaneous'). The concordance is 75.3 percent, which is statistically significant at the level of $p < .0005$. Thus, in three out of four patients, the project and computer diagnoses agree, which is a demonstration of the lack of 'randomness' in project diagnoses.

Conclusions

Psychiatric diagnoses made by standardized techniques can achieve a high degree of reliability as measured either by inter-rater comparisons of item ratings (92 percent) and diagnoses (80 percent) or by correlation with the computer diagnoses obtained by the CATEGO program (75 percent).

A comparison of standardized project consensus diagnoses with routine hospital diagnoses in three countries, revealed that the agreement between the two was highest in Toronto (80 percent), lowest in New York (59 percent), and intermediate in London (64 percent). This contrasts with the findings of the earlier US-UK videotape study, which concluded that Canadian psychiatrists were closer to their American colleagues than to their British counterparts in diagnostic habits, but it is quite possible that other Canadian centres may follow American diagnostic propensities. Further studies in other centres may resolve this issue.

In the psychogeriatric part of the Toronto sample, it was possible to compare project diagnoses in Toronto directly with those in New York and London, and there were no differences in morbidity between the three centres. It is likely that results would be similar in the younger age groups were a random sample selected from all hospital admissions for acute psychiatric illness.

These findings highlight the critical importance of controlling for diagnosis in psychiatric research, for example, American studies of metabolic debris in schizophrenic patients may include some who would not be so labeled in Canada, and

therefore the conclusions of such studies may not be true if applied directly to a Canadian patient population. Only by carefully standardized methods of interview and diagnosis can such pitfalls be avoided.

Summary

National statistics of mental illness suggest great differences in morbidity patterns between Canada, the USA and the UK, with the highest rates of illness occurring in Canada. Statistically, Canada and the US possess higher rates of schizophrenia and diseases of the senium but lower rates of affective illnesses than the UK. The US-UK study demonstrated that when the diagnostic process was standardized, the differences between the two countries disappeared. This paper examines a sample of 255 Canadian patients hospitalized in Toronto and compares the results of standardized techniques of diagnosis with the routine hospital diagnoses of the same patients.

Conclusions of the study were fourfold. First, psychiatric diagnoses made by standardized techniques can achieve a high degree of reliability as measured either by inter-rater comparisons of item ratings (92 percent) and diagnoses (80 percent) or by correlation with computer diagnoses obtained by the CATEGO program (75 percent). Second, comparisons of project consensus diagnoses with routine hospital diagnoses revealed that agreement between the two was highest in Toronto, lowest in New York and intermediate in London. Third, a comparison of the project diagnoses of psychogeriatric patients in the three centres revealed no differences in morbidity between the three countries. Finally, the fallacy of generalizing from studies in one country which have not controlled for diagnosis to populations in other countries, is exposed.

Acknowledgements

Grateful acknowledgment is given to the staffs of the Queen Street Mental Health Centre, the Toronto General Hospital and the Clarke Institute of Psychiatry for all their help and cooperation.

References

1. Cooper, J.E., Kendall, R.E., Gurland, B.J., Sharpe, L., Copeland, J.R.M., Simon, R.: *Psychiatric Diagnosis in New York and London*. Maudsley Monograph #20, 1972.
2. Copeland, J., Kelleher, M., Kellett, J.M., Gourlay, A.J., Cowan, D.W., Barron, G., De Gruchy, J. — UK Team. Gurland, B.J., Sharpe, L., Simon, R., Kuriansky, J. — US Team. "Cross-national study of diagnosis of mental disorders". *Br. J. Psychiatry*, 126, 11-20, 1975.
3. D.B.S. Mental Health Statistics, Vol. 1. 1969.
4. Duckworth, G.S., Ross, H.: Diagnostic differences in psychogeriatric patients in Toronto, New York and London, England. *C.M.A. J.* 112, p. 47, 1975.
5. Kedward, H.B., Eastwood, M.R., Allodi F., Duckworth, G.S.: The evaluation of chronic psychiatric care. *C.M.A.J.* 110, 519, 1974.
6. Kraepelin, E.: *Clinical Psychiatry*. Transl. A.R. Defendorf, Toronto, MacMillan Co., p. 201, 1904.
7. Neumann, H.: *Lehrbuch der Psychiatrie*. Erlangen, Germany: p. 187, 1859.
8. Sharpe, L., Gurland, B.J., Fleiss, J.L.: Comparisons of American, Canadian and British psychiatrists in their diagnostic concepts. *Can. Psychiatr. Assoc. J.* 19, p. 235, 1974.
9. Spitzer R.L., Fleiss, J.L., Burdock, E.I., Hardesty, A.S.: The mental status schedule: Rationale, reliability and validity *Compr. Psychiatry*, 5: 6, 384, 1964.
10. Wing, J.K. "A Standard Form of Psychiatric Present State Examination" In: *An International Symposium: Psychiatric Epidemiology*. Eds. Hare E.H. and Wing J.K.: Oxford University Press, England, 1970.

Résumé

Les statistiques nationales sur les maladies mentales présentent de grandes différences dans les types de morbidité, au Canada, aux Etats-Unis et en Grande-Bretagne. C'est au Canada que se trouve les taux les plus élevés de maladies mentales; la schizophrénie et les psychoses séniles sont plus fréquentes et les psychoses affectives moins nombreuses au Canada et aux Etats-Unis qu'en Grande-Bretagne. Dans l'étude portant sur les Etats-Unis et la Grande-Bretagne, les différences s'aplanissent après une formalisation des méthodes de diagnostic. L'auteur étudie un échantillon de 255 malades canadiens hospitalisés à Toronto et compare les résultats des techniques standardisées de diagnostic avec les diagnostics hospitaliers de routine posés pour la même population.

Quatre conclusions se dégagent de cette étude. D'abord, les diagnostics psychiatriques faits selon des techniques standardisées peuvent atteindre un haut degré de fidélité; soit qu'on utilise les comparaisons inter-juges dans les évaluations d'éléments (92%) et dans les diagnostics (80%), soit que l'on utilise une corrélation avec les diagnostics obtenus par ordinateur, suivant la méthode Catego (75%). Ensuite, les comparaisons des diagnostics collectifs de l'étude avec les ordinaires de l'hôpital démontrent une concordance la plus élevée à Toronto, la plus faible à New-York et, entre les deux, à Londres. Les diagnostics posés au cours de cette recherche chez les malades gériatriques ne présentaient aucune différence de la morbidité entre les trois pays. Enfin, l'auteur insiste sur l'erreur et le danger que comporte toute généralisation d'études diagnostiques faites dans un pays sans contrôle à d'autres populations.